

PCT

THE PATENT LITRARY  
SCIENCE REFERENCE AND INFORMATION SERVICE

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



B

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12Q 1/68</b>		A1	(11) International Publication Number: <b>WO 98/42870</b>
			(43) International Publication Date: 1 October 1998 (01.10.98)
(21) International Application Number: <b>PCT/US98/05550</b>		(81) Designated States: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CU, CZ, EE, GE, GW, HU, ID, IL, IS, JP, KG, KP, KR, KZ, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, SL, TJ, TM, TR, TT, UA, US, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 19 March 1998 (19.03.98)			
(30) Priority Data: 60/041,515 24 March 1997 (24.03.97) US			
(71) Applicant (for all designated States except US): E.I. DU PONT DE NEMOURS & COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).			
(72) Inventor; and			
(75) Inventor/Applicant (for US only): REITER, Robert, Stefan [US/US]; 4334 101st Street, Urbandale, IA 50322 (US).			
(74) Agent: MAJARIAN, William, R.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).			
<b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>			
(54) Title: A METHOD TO IDENTIFY AND BREED CORN WITH INCREASED KERNEL OIL CONCENTRATION			
(57) Abstract <p>A method for breeding with high oil corn germplasm is disclosed. The method involves the use of genetic markers associated with trait loci controlling kernel oil concentration. These genetic markers are used to select for kernel oil concentration in breeding populations. Also disclosed is a method for selecting complementary oil parent sources using genetic markers, which are likely to produce superior offspring. Also disclosed are the trait loci controlling corn kernel oil concentration.</p> <p><i>10/5/98</i> <i>W. E. Miller</i></p>			

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

# A METHOD TO IDENTIFY AND BREED CORN WITH INCREASED KERNEL OIL CONCENTRATION

5       The invention is in the fields of plant breeding and molecular biology. More specifically, the invention relates to the identification of corn loci conferring increased kernel oil concentration using genetic markers and the use of genetic markers as an aid to the identification and breeding of corn with increased kernel oil concentration.

Corn is a major crop used as a human food source, an animal feed, and as a source of carbohydrate, oil, protein, and fiber. It is principally used as an energy source in animal feeds, or as a raw material for the recovery of starch, protein feed fractions, fiber, flaking grits, flour, and oil.

The heritability of a trait is defined in the broad sense as the ratio of the genetic variance to the total phenotypic variance. Many agronomic traits display low heritability; i.e., the performance of parent plants is a poor predictor of offspring performance. Thus, traits with low heritability have small genetic variance components in comparison with observed variation. The impact on the plant breeder is that in breeding populations, the value of a plant's genetic composition is difficult to determine from agronomic trait measurements. In an attempt to maximize their discriminative abilities, breeders collect multiple measurements both from individuals related by descent and from many environments. This strategy is resource intensive because it involves the use of extensive trialing to make even small gains in plant improvement. This, coupled with the fact that improved corn lines are selected for multiple traits simultaneously, makes the development of superior corn inbreds both a time-consuming and an expensive labor.

The addition of novel traits in a corn breeding program imposes an additional burden upon the plant breeder. Depending upon the genetic complexity of the novel trait (i.e., single gene versus many genes), a significant increase in time and effort is required to produce elite lines containing novel traits. One such

5 trait is kernel oil concentration.

Corn with increased kernel oil concentration is important because it possesses improved feeding value for poultry (Han Y. et al. (1987) *Poultry Sci.* 66:103-111) and livestock (Nordstrom, J.W. et al. (1972) *J. An. Sci.* 35(2):357-361). Grain from conventional corn hybrids typically contains 4% oil.

10 In an effort to increase the kernel oil concentration, a long-term recurrent selection program was initiated in the open-pollinated cv. Burr's White by C.G. Hopkins in 1896. This recurrently-selected population known as Illinois High Oil (IHO), has been selected for increased oil concentration for over ninety generations (Dudley, J.W. and R.J. Lambert. (1992) *Maydica* 37:1-7) using modified mass selection.

15 As a result, oil concentration was increased in the population over 20%. The germplasm was little used because derived materials had yields substantially lower than conventional varieties (Alexander, D.E. (1988) In: Proc. 43rd Ann. Corn and Sorghum Res. Conf. Am. Seed Trade Assoc., Washington, D.C. pp 97-105).

20 Using thirty-eight open-pollinated cultivars and synthetics, Alexander initiated a second recurrent selection program (Alexho synthetic) to increase kernel oil (Alexander, D.E. (1988) In: Corn and Corn Improvement. G.F. Sprague and J.W. Dudley eds. American Society of Agronomy, Madison WI. Pp 869-880). Equivalent oil levels to IHO were achieved in twenty-eight cycles using selection

25 based upon the oil concentration of single ears and in later generations based upon the oil concentration of single kernels. Yield performance of Alexho-derived material in single cross hybrids (high oil inbred x conventional inbred) is improved over IHO, presumably due to the greater genetic variability initially available, although performance was not equivalent to conventional hybrids. The

30 development of agronomically elite corn germplasm also containing increased kernel oil concentration is clearly a challenge using conventional plant breeding methods.

Kernel oil concentration can be phenotypically measured using a variety of analytical methods. Oil concentration displays a non-discrete distribution,

35 common for quantitatively-inherited traits controlled by several loci. Kernel oil measurements select those breeding lines with the highest phenotypic expression. Unfortunately, the genetic potential for high oil is limited in most of these lines because it is impossible to discriminate between lines based upon their true genetic composition. This situation is further aggravated when simultaneous

selection for agronomic performance is practiced. It would therefore be advantageous to base selection upon the genotype of the plants in the population. Genetic markers, especially nucleic acid markers, may be used to advantage as an indirect selection method for complex quantitative traits. Genetic markers  
5 identifying alleles conferring increased oil would therefore be an advantageous tool for plant breeding programs developing elite high oil corn germplasm.

There is limited published information on the identification of genetic markers predictive for increased oil yield. Kahler (Kahler, A.L. (1985) In: Proc. 40th Ann. Corn and Sorghum Res. Conf. Am. Seed Trade Assoc., Washington  
10 D.C. pp. 66-89) measured isozyme allelic frequency changes following twenty-five cycles of selection in Alexho synthetic and found eight significant loci. Most of these allele frequency changes were also significant for tests measuring random genetic drift, making it difficult to conclude that selection based upon these isozyme alleles would be useful. More recently Goldman et al. (Goldman, I.L., et  
15 al. (1994) *Crop Sci.* 34:908-915) and Berke and Rocheford (Berke, T.G. and Rocheford, T.R. (1995) *Crop Sci.* 35:1542-1549) used RFLP markers to identify significant marker loci associated with oil concentration in the Illinois long-term selection populations. These studies identified twenty-five and thirty-one markers respectively, in populations derived from Burr's White, which were significantly  
20 associated with increased oil. Some of the regions identified by significant RFLP marker loci may be in common between the two studies, however of the fifteen RFLP markers which were used in both studies, six were in disagreement for their effect on oil concentration. In these studies the populations used were derived from common ancestry (Burr's White); however, the populations were selected for  
25 different traits (oil and protein) over many generations. It is not surprising that many identified oil loci would be unique to each population analyzed. It is therefore desirable to identify those genetic markers which are uniquely predictive of germplasm being used in the breeding program.

#### SUMMARY OF INVENTION

30 A method is disclosed for reliably and predictably breeding for corn with increased kernel oil concentration. The method comprises a) using one or more genetic markers to select a corn plant from a corn breeding population by marker-assisted selection, wherein the genetic markers are selected from the group consisting of s1375, s1384, s1394, s1416, s1422, s1432, s1457, s1480, s1476,  
35 s1478, s1484, s1500, s1513, s1529, s1544, s1545, s1630, s1633, s1647, s1750, s1756, s1757, s1767, s1772, s1774, s1780, s1797, s1813, s1816, s1817, s1836, s1853, s1860, s1870, s1921, s1922, s1925, s1931, s1933, s1939, s1946, s1949, s2054, s2055, s2057, s2058, s2097, s2122, s2125, s2150, s2156 and s2175; and b) crossing the selected corn plant with a second corn plant wherein the progeny

of the cross displays increased kernel oil concentration. A preferred source of high oil corn germplasm is a member of an Alexho synthetic population or a progeny thereof.

Also disclosed is a method for identifying corn plants or corn lines for use as parents for creation of a breeding population, the method comprising

5 a) genotyping corn plants or corn lines with one or more genetic markers wherein the genetic markers are selected from the group consisting of s1375, s1384, s1394, s1416, s1422, s1432, s1457, s1480, s1476, s1478, s1484, s1500, s1513, s1529, s1544, s1545, s1630, s1633, s1647, s1750, s1756, s1757, s1767, s1772, s1774,

10 s1780, s1797, s1813, s1816, s1817, s1836, s1853, s1860, s1870, s1921, s1922, s1925, s1931, s1933, s1939, s1946, s1949, s2054, s2055, s2057, s2058, s2097, s2122, s2125, s2150, s2156 and s2175; and b) identifying corn plants or corn lines which, based upon their genotype, are predicted to produce transgressive segregants for kernel oil concentration.

15 The present invention provides a method for the identification of and selection for genes controlling increased corn kernel oil concentration. These oil alleles were initially identified in materials composed of or derived from the Alexho synthetic breeding populations. Further, the method facilitates the use of this high oil material in breeding programs with the objective of developing new

20 high oil corn germplasm.

Specifically, the method uses genetic markers to predict the oil breeding value of lines in a corn breeding program. By indirect selection of oil loci using these markers, those lines with the greatest genetic potential for increased kernel oil concentration are chosen.

25 According to the method, any type of genetic marker may be used to identify an association with kernel oil concentration. The method is only limited by the ability to measure polymorphism at a given marker locus. Those skilled in the art will recognize that the various genetic markers which may be used includes but is not limited to restriction fragment length polymorphisms (RFLPs), random

30 amplified polymorphic DNAs (RAPDs), simple sequence repeats (SSRs), AFLPs, various single base pair detection methods, allozymes, and phenotypic markers. SSR markers useful in the practice of the instant method include s1375, s1384, s1394, s1416, s1422, s1432, s1457, s1476, s1478, s1480, s1484, s1500, s1513, s1529, s1544, s1545, s1630, s1633, s1647, s1750, s1756, s1757, s1767, s1772,

35 s1774, s1780, s1797, s1813, s1816, s1817, s1836, s1853, s1860, s1870, s1921, s1922, s1925, s1931, s1933, s1939, s1946, s1949, s2054, s2055, s2057, s2058, s2097, s2122, s2125, s2150, s2156 and s2175.

A further embodiment of the present invention are the trait loci controlling the expression of corn kernel oil concentration. These loci are identified and defined (i.e., mapped) by the marker loci of the present invention.

- 5 An additional embodiment of the present invention are corn plants and high oil corn germplasm that are produced using the instant breeding method.

#### DETAILED DESCRIPTION OF THE INVENTION

- 10 Table 1 provides a brief description of the genetic markers that form a part of the instant invention. Each marker is defined by its constituent nucleic acid primers (forward and reverse) that facilitate amplification of the specific marker locus of the corn genome. Also indicated is the required identifier for each sequence. The identifiers listed in Table 1 correspond to those listed in the Sequence Listing (*infra*) as required by 37 C.F.R. §1.821 et seq.

15 Table 1  
Genetic markers useful for defining the location of trait loci  
controlling corn kernel oil concentration

Marker	Sequence (5'-3')	Primer Type	SEQ ID NO.
s1375	TTTATGGGTTGGGAGATACTTG	forward	1
	AGATGTGTGCGTTTTTGAGAG	reverse	2
s1384	TTACGGCCTAGACATTTTCGAC	forward	3
	CACTTGCTTTCAGGTACCCA	reverse	4
s1394	CTGCCCAGTCCGTAATGAA	forward	5
	TAGATTTATTTTCTGAACGATTGG	reverse	6
s1416	GATCTCTCTGAGGCTTGTC	forward	7
	TGTAGTTGAGGATGCTCCC	reverse	8
s1422	AGGCAAGGCTTCTTCATAC	forward	9
	CGGACGACGACTGTGTTT	reverse	10
s1432	ACATGAGAAACAAGATAGAACCAG	forward	11
	AAAATGTAAGAACTTGTTTGGGA	reverse	12
s1457	CTGCTTATTGCTTTCGTCATA	forward	13
	TGCTGCACTACTTGAACCTAG	reverse	14
s1476	ACACAGAGATGACAAAAGCAA	forward	15
	GCAGGCGTGCTATGAGAG	reverse	16
s1478	AGCGGTGAAACCCTTATG	forward	17
	CTGTGGCTGGTTCCTCTC	reverse	18
s1480	GCTCTTGATAAAAAGGCAAGT	forward	19
	CTTGTTGTAATGGATGAGTGAG	reverse	20
s1484	GCTCGTAGTAGGGGTTACG	forward	21
	GACAGCCTCACCTCAAGA	reverse	22

s1500	ACAGATCTTGACACGTACATACC	forward	23
	GGACGTGTATCCTCAAATCAT	reverse	24
s1513	CAGCGAATACTGAATAACGC	forward	25
	TGTTGGATGAGCACTGAAC	reverse	26
s1529	TGTTCTCAACAACCACCG	forward	27
	CGTTTAGCGATATCATTTTCC	reverse	28
s1544	GATCCTACCAAAATCTTATAGGC	forward	29
	ACAGCTAGCCAAGATCTGATT	reverse	30
s1545	CGATACTAATGGAAGCCCTAA	forward	31
	ATGGCCCATTAAGTTTATCAC	reverse	32
s1630	AAAGCGTAGTCGGAAAGC	forward	33
	ACCAATGATCTTTACGCAGAT	reverse	34
s1633	TAATCAGAGCGTACATCAGGA	forward	35
	AGGGCATCAATCAAGAATG	reverse	36
s1647	GAGACTTTTGAGGAGAAAGCA	forward	37
	GATCAAAAGAGCAAAAGGAGA	reverse	38
s1750	AACTGATGAATACCTTCCCAG	forward	39
	TGATTAACCTTCTCCCTTTGGT	reverse	40
s1756	TCGGCACAACATATGAGTTAC	forward	41
	CCCCCATAGAGAGAGATAGAG	reverse	42
s1757	AAGCACGGCCCAATAGAAT	forward	43
	AGGATGTCCCTAGCTTTATTG	reverse	44
s1767	TCATTGCCCAAAGTGTTG	forward	45
	CTCATCACCCCTCCAGAG	reverse	46
s1772	GATCCACGCCATTTAAAC	forward	47
	TGATACTCTGGTGCATGTTT	reverse	48
s1774	GATCGCTCCGATCTATCC	forward	49
	AGCGGCATCTATGTTCTATG	reverse	50
s1780	CCCAGTGCGAAGAGACTC	forward	51
	ACACCTGCTCTGCACCAC	reverse	52
s1797	CTAACCACGACGACCCT	forward	53
	GCATGAGTGCATGTGCAT	reverse	54
s1813	CTGCCACATGCTTTTCTG	forward	55
	CTGTAAAGAAGCTGGTCTGGA	reverse	56
s1816	TTCTCCTCATGGATGCGT	forward	57
	CTATTTGGAAGTATGGGCTTCA	reverse	58

s1817	GAGGGCATCTATGTGCAAC	forward	59
	GCTCAGAAAGTTGCGTTTATG	reverse	60
s1836	TTCCTTCACGTTTCTCTGTAA	forward	61
	CACATAAACCTAATGGGGTACA	reverse	62
s1853	CCCAAAGGCGATACCTATT	forward	63
	CCCACTTTCTCACTCTTTTCT	reverse	64
s1860	GAGGTGAGTACTATGCAAATGC	forward	65
	CAGGCTTACCTAGCCTTCTC	reverse	66
s1870	CTATGGATGGCTGCTTGC	forward	67
	GTCAGGCAGCAGAATGTG	reverse	68
s1921	AAACCGTCCAGCGACTAC	forward	69
	GGAAGAACCAATCCCATATCT	reverse	70
s1922	AACATCCTGTCGGAAACAG	forward	71
	TCATCACGTCTCTCTTTCAAC	reverse	72
s1925	TTGTGGCAGAATCTCAAATTA	forward	73
	CGACTGGTGACATGTGAAG	reverse	74
s1931	AGTGAGGAAAGAATATGCTGG	forward	75
	TGGACTGAGAACTGATTTGA	reverse	76
s1933	CACAAATGTGAAGGTAAACACT	forward	77
	AATGGTACGGTTCAGGATG	reverse	78
s1939	AGATGACGCACGGAACAC	forward	79
	AGCATCATGTAGCAGGAGG	reverse	80
s1946	TTGCAGCACTGTCGTAGTC	forward	81
	GCGCGAGTGGAGTAGTAAG	reverse	82
s1949	AAGATTATGCAGATGAGACACC	forward	83
	GTTCCATGCTTTCCTTGG	reverse	84
s2054	GCCGATACCATGTAAGAGAAT	forward	85
	CTCTGGGCTCTGTGTTAGAGT	reverse	86
s2055	CTGCTTTCTCTGTTCCAGC	forward	87
	AATCGCTTACTTGTAACCCAC	reverse	88
s2057	AAGAACGTACGTCCCATAAAG	forward	89
	CAAGGTAAAGTGACAAAGCAG	reverse	90
s2058	G TTCAGGATGAGGCGGAA	forward	91
	GTGATCATCGCAGGAGACC	reverse	92
s2097	GGAGCCTGGAGTGAGAAC	forward	93
	CATGCTCACCTAACGTGG	reverse	94

s2122	ATCTGAACACTTGAGCAACAA	forward	95
	ATAGACCGGACCCATCAC	reverse	96
s2125	CGAACAGCGGGTACACCT	forward	97
	GAGGTCAGCTTCCTCGATCT	reverse	98
s2150	GGAATCGTTCCTCCACAC	forward	99
	CTTCCTCGGTGTCAGACG	reverse	100
s2156	ATGGAAACATCAAAGTGGATT	forward	101
	TGCTACCCTGATGACCTGAT	reverse	102
s2175	ACCACTAGTCTCATATGAAGGG	forward	103
	GGTAGGTGGGTAGGGGTT	reverse	104

For the purposes of this invention, we define the following terms:

Corn. Any variety, cultivar, or population of Zea mays L.

5 Elite. This term characterizes a plant or variety possessing favorable traits, such as, but not limited to high yield, good grain quality, and disease resistance. This enables its use in commercial production of seed or grain at a profit. The term also characterizes parents giving rise to such plants or varieties.

10 High Oil Corn Germplasm. This term characterizes corn plants which, when either self-pollinated or used as either the male or the female parent in a variety of outcrossing combinations, produce kernels with increased oil when compared to kernels produced by non-high oil germplasm. Examples of high oil corn germplasm include but are not limited to open-pollinated varieties, hybrids, synthetics, inbred lines, races, and populations or corn plants derived from one of

15 the aforementioned.

Variety or cultivar. These terms refer to a group of similar plants that by structural features and performance can be identified from other varieties or cultivars within the same species.

20 Line. This term refers to a group of individuals from a common ancestry; a more narrowly defined group than a variety.

Synthetic. This term refers to a genetically heterogeneous collection of plants of known ancestry created by the intermating of any combination of inbreds, hybrids, varieties, populations, races or other synthetics.

25 Inbred. This term refers to a substantially homozygous individual, variety or line.

Recombinant Inbreds. A population of independently derived lines developed by repeated selfing each generation until complete homozygosity is approached. Each recombinant inbred is derived from a single F2 plant using a breeding method commonly referred to as single seed descent.

Breeding. The art and science of improving a species of plant or animal through controlled genetic manipulation.

Marker-Assisted Selection. The use of genetic markers to identify and select plants with superior phenotypic potential. Genetic marker(s) determined previously to be associated with a trait locus or trait loci are used to uncover the genotype at trait loci by virtue of linkage between the marker locus and the trait locus. Plants containing desired trait alleles are chosen based upon their genotypes at linked marker loci.

Alexho Synthetic. Recurrently selected, high oil corn germplasm developed by Denton Alexander at the University of Illinois. Alexho synthetic high oil corn germplasm is composed of multiple synthetic populations defined by their cycle of advancement in the recurrent selection breeding program.

Breeding Population. A genetically heterogeneous collection of plants created for the purpose of identifying one or more individuals with desired phenotypic characteristics.

Phenotype. The observed expression of one or more plant characteristics.

Phenotypic Value. A measure of the expected expression of an allele at a trait locus. The phenotypic value of an allele at a trait locus is dependent upon its expressive strength in comparison to alternative alleles. The phenotypic value of an individual, and hence its phenotypic potential, is based upon its total genotypic composition at all loci for a given trait.

Transgressive Segregants. Individuals whose phenotype exceeds the phenotypic variation predicted by the parents.

Genetic Marker. Any morphological, biochemical, or nucleic acid based phenotypic difference which reveals a DNA polymorphism. Examples of genetic markers includes but is not limited to RFLPs, RAPDs, allozymes, SSRs, and AFLPs.

Marker locus. The genetically defined location of DNA polymorphisms as revealed by a genetic marker.

Trait Locus. A genetically defined location for a collection of one or more genes (alleles) which contribute to an observed characteristic.

Genotype. The allelic composition of an individual at genetic loci under study.

Restriction Fragment Length Polymorphism (RFLP). A DNA-based genetic marker in which size differences in restriction endonuclease generated DNA fragments are observed via hybridization (Botstein, D. et al. 1980. *Am. J. Hum. Genet.* 32: 314-331).

Random Amplified Polymorphic DNA (RAPD). A DNA amplification-based genetic marker in which short, sequence arbitrary primers are used and the

resulting amplification products are size separated and differences in amplification patterns observed (Williams J.G.K. et al. 1990. *Nucleic Acids Res.* 18:6531-6535).

Simple Sequence Repeat (SSR). A DNA amplification-based genetic marker in which short stretches of tandemly repeated sequence motifs are amplified and the resulting amplification products are size separated and differences in length of the nucleotide repeat are observed (Tautz D. 1989. *Nucleic Acids Res.* 17:4127-4138).

AFLP. A DNA amplification-based genetic marker in which restriction endonuclease generated DNA fragments are ligated to short DNA fragments which facilitate the amplification of the restricted DNA fragments (Vos, P. et al. 1995. *Nucleic Acids Res.* 23:4407-4414). The amplified fragments are size separated and differences in amplification patterns observed.

Allozymes. Enzyme variants which are electrophoretically separated and detected via staining for enzymatic activity (Stuber, C.W. and M.M. Goodman. 1983. USDA Agric. Res. Results, Southern Ser., No. 16).

The present invention relates to the discovery of trait loci controlling kernel oil concentration through the use of genetic markers. In populations in which variation for both kernel oil concentration and genetic marker alleles exist, oil measurements and marker-based genotypes were generated for members of the populations. Using least squares methods, the locations of oil concentration loci were determined in relation to markers genetically linked to these trait loci. Indirect selection of preferred oil alleles may now be practiced using the information at one or more linked genetic markers. Selected corn plants comprise one or more alleles encoding a high oil phenotype.

It is recognized that several different populations and population types could be used to locate trait loci of interest. Some of the population types include but are not limited to recombinant inbreds, backcrosses, F2's or their self-pollinated or intermated derivatives, and synthetics. Further, it is understood that an alternative to measuring phenotypic and genotypic variation within populations is the measurement of genotypes and phenotypes between populations. In this alternative the second population is a selected derivative of the first population, selection being either on the trait of interest (phenotypic selection) or on specific marker alleles (genotypic selection). It is also recognized by those skilled in the art that alternative statistical approaches may be used to determine a linkage relationship between marker loci and trait loci.

#### EXAMPLES

The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the

above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

5

#### EXAMPLE 1

### LOCATION OF LOCI CONFERRING INCREASED KERNEL OIL CONCENTRATION

#### Population development and trait measurement

LH119wx and LH51, two inbred corn lines developed by Holden's Foundation Seed Co., Williamsburg, IA were independently intermated with individual plants from the synthetic population ASKC28wx (deposited at the American Type Culture Collection, Rockville, MD; Accession No. ATCC 75105) (waxy kernels are highly represented in ASKC28 and as such I have designated the ASKC28 as being waxy). The F1 plants were selfed and resulting F2 populations were grown. Individual F2 plants were selfed and derived kernels were advanced using single seed decent through six generations of selfing (S6) to produce recombinant inbred lines. Up to twenty kernels from the S6 generation were grown and selfed producing a family of S7 ears representing each recombinant inbred line. Oil values were determined for each ear within a family using near infrared transmittance (Williams, P.C. (1987) In: Near Infrared Technology in the Agricultural and Food Industries; P.C. Williams and C. Norris, eds. American Association of Cereal Chemists).

#### Genotypic determination

Ten seeds from single ears representing each of one hundred ninety-four (LH119wx x ASKC28wx) or two hundred and four (LH51 x ASKC28wx) recombinant inbred lines were germinated on moistened filter paper. Root segments were excised from germinated seeds, pooled for each ear and extracted using an automated DNA extraction machine. The instrument uses a modification of the Murray and Thompson CTAB procedure (Murray, M.G. and Thompson, W.F. (1980) *Nucl. Acids Res.* 8:4321-4325). DNA samples were quantified via fluorescence using YoPro-1™ iodide (Molecular Probes, Inc., Eugene, OR) and diluted to 4 µg/ml.

SSR regions for each DNA sample were analyzed using the following protocol:

1. Ten µl of amplification cocktail (see Table 2) was added to 5 µl (20 ng) of extracted DNA;
2. The DNA fragment flanked by sequences complementary to the primers present in the amplification cocktail was amplified by PCR (U.S. Patent No. 4,683,202 and U.S. Patent No. 4,683,195) using the following protocol:

1) 45 cycles of 50 sec at 95°C, 50 sec at 54°C and 80 sec at 72°C and 2) 1 cycle of 300 sec at 72°C;

3. Approximately 8 µl of each sample was loaded onto agarose gels composed of 2% Metaphor (FMC Corp., Rockland, ME), 1X TBE, and 0.5 µg/ml ethidium bromide, and electrophoresed for 2 h at 6.1 V/cm in horizontal electrophoresis units to which 1X TBE buffer and 0.5 µg/ml ethidium bromide was added; and

4. DNA bands were visualized by UV fluorescence.

10

Table 2  
Amplification Cocktail

Reagent	Stock Concentration	Final Concentration
Buffer*	10X	1.5X
dNTPs	2 mM	0.3 mM
Forward Primer	40 µM	0.45 µM
Reverse Primer	40 µM	0.45 µM
AmpliTaq Polymerase™	5 U/µl	0.05 U/µl

\*10X Buffer is a pH 9.0 solution composed of 800 mM Tris-OH, 200 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 25 mM MgCl<sub>2</sub>.

#### 15 Localization of oil loci

One hundred thirty three polymorphic SSR marker loci were used to genotype the recombinant inbreds from the LH119wx x ASKC28wx cross and one hundred and three polymorphic SSR marker loci were used to genotype the LH51 x ASKC28wx-derived population. In addition, twenty publicly available polymorphic SSR loci with previously established chromosome locations and covering all ten maize chromosomes (available from Research Genetics, Huntsville, AL) were also mapped in both populations.

Genetic linkage and distance between marker loci was determined independently for each population using MAPMAKER 3.0 (Lincoln S.E., et al. (1993) Whitehead Inst. Biomed. Res., Cambridge, MA). This resulted in the establishment of ten linkage groups for each population corresponding to the ten chromosomes of maize. Each linkage group was assigned to a chromosome based upon linkage to the public SSR markers. Twenty-three and ten markers in the LH119wx x ASKC28wx and LH51 x ASKC28wx populations, respectively, were not assigned chromosome positions because genetic linkage could not be clearly established.

Analysis of variance was used to identify marker loci in linkage with trait loci conferring increased oil concentration. Oil concentration was used as a

dependent variable and separate ANOVAs were calculated with SAS Proc GLM (SAS Inst., Cary, NC) using each marker locus as a single independent variable (Edwards, M.D., et al. (1987) *Genetics* 116:113-125). Therefore, for each ANOVA test the mean oil values of marker allele classes were compared. Marker  
 5 loci were declared significant if  $p < 0.05$ .

Linkage data for significant marker loci was examined to determine both the number of trait loci present and their probable location. Significant marker loci on the same linkage group are either detecting the same trait locus or alternatively different trait loci. By careful examination of the phenotypic  
 10 variation explained by each marker locus along the chromosome, a determination of the number trait loci on a linkage group was made. Significant marker loci, on the same linkage group and uninterrupted by non-significant marker loci, were declared to be detecting the same trait locus on the chromosome. If significant marker loci on the same chromosome were interrupted by non-significant marker  
 15 loci then each significant region was declared to contain a trait locus resulting in multiple trait loci on the same chromosome.

To confirm the number of trait loci, marker data assigned to linkage groups and oil data were also analyzed with Mapmaker/QTL 1.0 (Lincoln, S.E. et al. (1990) Whitehead Inst. Biomed. Res., Cambridge, MA). Results with  
 20 Mapmaker/QTL were in agreement with the initial analysis for the number trait loci on each chromosome.

Eleven and twelve loci controlling kernel oil concentration were located in the LH119wx x ASKC28wx and LH51 x ASKC28wx recombinant inbred populations, respectively. Each oil locus is defined by one or more linked marker  
 25 loci.

In instances where the same marker loci were used in both populations, alignment of linkage groups is possible. It was found that in most instances both populations localized the same oil loci. By considering common marker loci, a total of seventeen loci controlling kernel oil concentration were found. Each oil  
 30 locus was assigned an arbitrary letter designation (Table 3).

Table 3

Marker loci genetically linked to and predictive of the location of trait loci  
conferring increased kernel oil concentration

Oil locus	Chromosome	Marker loci
A	1	s1922
B	1	s1478, s1853, s1949
C	1	S1860, s1925, s1931, s2150
D	2	s2175

E	3	s1394
F	4	s1476, s1772, s1816, s2122, s1836
G	4	s1939, s1946
H	4	s1870
I	5	s1529
J	5	s2054, s1647, s1500, s1545, s1774, s2097
K	6	s1457, s2055, s1757, s2125, s1780, s1375, s1797, s1416, s1432, s1921
L	7	s1630, s1422, s2156
M	8	s1817, s2057
N	9	s1544, s1633, s1384, s1813, s1767, s2058, s1933, s1513, s1484
O	10	s1756
P	10	s1480 (positive oil allele in LH51)
Q	N.A.*	s1750

\*N.A. - chromosome location not known

In instances where comparisons could be made, oil loci which were identified in one population were identified at the same location in the second population. In two exceptions, an oil locus was found in one population, but not in the second population. In the first case, the allele with a positive oil effect was found in LH51 and thus it would be unexpected to identify the same locus in the LH119wx x ASKC28wx population. In the second case, it was found that different ASKC28wx-derived marker alleles were segregating in the populations; therefore, each population was measuring the oil effect of a different ASKC28wx allele at the trait locus. The most abundant ASKC28wx oil allele segregating in LH119wx x ASKC28wx had a positive oil effect versus the alternative LH119-derived allele, whereas in the LH51 x ASKC28wx population, the abundant ASKC28wx allele had no positive oil effect. With the exception of the oil locus linked to marker s1480, all alleles with positive effects on oil concentration were derived from ASKC28wx.

#### EXAMPLE 2

#### MARKER-ASSISTED SELECTION OF BREEDING LINES USING GENETIC MARKERS FOR INCREASE KERNEL OIL CONCENTRATION

Genetic marker loci in linkage with oil trait loci are highly predictive of oil concentration and as such may be used as an indirect measurement of kernel oil in a marker-assisted selection program. Accordingly, genotypic information from linked marker loci would facilitate the selection of breeding lines with increased oil concentration. Direct oil measurements cannot differentiate between various

genotypic trait locus compositions with equivalent phenotypic effects. This is especially problematic in early generation segregating breeding populations where only limited fixation of oil loci has occurred.

By way of example, an objective of a corn breeding program could be the creation of new elite inbred lines which contain trait alleles conferring increased kernel oil concentration. These trait alleles would be introduced by the intermating of high oil germplasm with one or more elite corn inbreds. The resultant hybrid could be self-pollinated to produce an F<sub>2</sub> population for the purposes of initiating a conventional pedigree breeding program (Allard, R.W. (1960) Principles of Plant Breeding. John Wiley & Sons, Inc. New York. Pp 115-128).

In order to identify those F<sub>2</sub> individuals with the desired genotypes, plant tissue would be collected from each F<sub>2</sub> individual in the population and genotyped with the SSR marker loci listed in Table 1. Those F<sub>2</sub> individuals with the highest frequency of SSR marker alleles derived from the high oil source would be selected and further culled based upon their agronomic fitness. With continued inbreeding and segregation, those oil loci in a heterozygous state could become fixed for either the high oil or low oil allele. It is therefore likely that genotyping and selection of later generation materials would be practiced in order to further segregate breeding lines based upon their marker allele and hence oil allele composition.

Depending upon population size and serendipity, the resulting inbreds from the pedigree breeding program may not demonstrate sufficient agronomic competitiveness or sufficient kernel oil expression because an inadequate number of oil alleles was recovered. These new inbreds could therefore be used as parental material and new breeding projects initiated. The SSR markers could again be used for further selection of oil as described.

It is obvious to those skilled in the art that many variants to selection methodology may be envisioned. Selection would be based upon the allelic composition of one or more marker loci which identify trait oil loci present in a population. Further selection would be performed by examination and selection of genotypes from individual plants, families, or their progeny. Various predictive models could be developed using genotypic information, which could generate various selection indices. These models would permit weighting the effect predicted by marker loci. This is because the predictive value of an individual marker locus is dependent upon its genetic distance from the corresponding trait locus as well as the expressivity of the trait locus. Selection strategies which combine phenotype-based and genotype-based selection may also be envisioned.

The marker loci presented here are predictive of oil loci in Alexho synthetic populations. Because ASKC28wx represents the 28th oil breeding cycle of a genetically closed population, earlier breeding cycles are composed of the same oil loci. It is expected that cycles differ simply in their allelic frequency at the identified oil loci. Therefore, in breeding populations derived from earlier Alexho cycles, the marker loci described in this invention will be useful in identification of oil loci and in prediction of oil concentration.

### EXAMPLE 3

#### IDENTIFICATION OF CORN PLANTS FOR USE AS PARENTS FOR THE PRODUCTION OF TRANSGRESSIVE SEGREGANTS FOR KERNEL OIL CONCENTRATION

It is important to identify corn plants and lines which, when used as parents, have the greatest probability of producing offspring with superior performance. Transgressive segregant offspring of such parents would result from the crossing of parents with complementary sets of alleles conferring the high-oil phenotype. Using the information provided herein, marker alleles which predict desired trait performance (i.e., high oil) at a given marker locus are known. By genotyping lines at those marker loci, the value of those lines as parents is revealed. For example, if one wanted to create an individual containing superior alleles at 5 separate oil loci (A-E), one could identify and cross a parent composed of desired alleles for locus A, B, and C with a parent composed of desired alleles at B, D, and E. These parents are complementary because they permit the recovery of progeny containing desired alleles at all 5 loci. Ideally, parents would be chosen which when combined ensure maximum complementation of loci, so that a high frequency of desired recombinants are recovered.

SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT:
  - (A) ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
  - (B) STREET: 1007 MARKET STREET
  - (C) CITY: WILMINGTON
  - (D) STATE: DELAWARE
  - (E) COUNTRY: USA
  - (F) ZIP: 19898
  - (G) TELEPHONE: 302-992-4926
  - (H) TELEFAX: 302-773-0164
  - (I) TELEX: 6717325
- (ii) TITLE OF INVENTION: A METHOD TO IDENTIFY AND BREED  
CORN WITH INCREASED KERNEL OIL  
CONCENTRATION
- (iii) NUMBER OF SEQUENCES: 104
- (iv) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE DISKETTE, 3.50 INCH
  - (B) COMPUTER: IBM PC COMPATIBLE
  - (C) OPERATING SYSTEM: MICROSOFT WINDOWS 95
  - (D) SOFTWARE: MICROSOFT WORD VERSION 7.0A
- (v) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vi) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 60/041,515
  - (B) FILING DATE: MARCH 24, 1997
  - (C) CLASSIFICATION:
- (vii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: MAJARIAN, WILLIAM R.
  - (B) REGISTRATION NUMBER: P-41,173
  - (C) REFERENCE/DOCKET NUMBER: BB-1076

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTTATGGGTT GGGAGATACT TG

22

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AGATGTGTGC GTTTTTGAGA G

21

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTACGGCCTA GACATTCGA C

21

## (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CACTTGCTTT CAGGTACCCA

20

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTGCCCAGTC CGTAATGAA

19

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TAGATTTATT TTCTGAACGA TTGG

24

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GATCTCTCTG AGGCTTGTCC

20

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TGTAGTTGAG GATGCTCC

19

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGCAAGGCT TTCTTCATAC

20

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGGACGACGA CTGTGTTC

18

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACATGAGAAA CAAGATAGAA CCAAG

24

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AAAATGTAAG AACTTGTTTG GGA

23

## (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTGCTTATTG CTTTCGTCAT A

21

## (2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGCTGCACTA CTTGAACCTA G

21

## (2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ACACAGAGAT GACAAAAGCA A

21

## (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAGGCGTGC TATGAGAG

18

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGCGGTGAAA CCCTTATG

18

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTGTGGCTGG TTCCTCTC

18

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCTCTTGATA AAAAGGCAAG T

21

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTTGTTGTAA TGGATGAGTG AG

22

## (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCTCGTAGTA GGGGTTACG

19

## (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GACAGCCTCA CCTCAAGA

18

## (2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ACAGATCTTG ACACGTACAT ACC

23

## (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGACGTGTAT CCTCAAATCA T

21

## (2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CAGCGAATAC TGAATAACGC

20

## (2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TGTTGGATGA GCACTGAAC

19

## (2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TGTTCTCAAC AACCACCG

18

## (2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CGTTTAGCGA TATCATTTTC C

21

## (2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GATCCTACCA AAATCTTATA GGC

23

## (2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ACAGCTAGCC AAGATCTGAT T

21

## (2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGATACTAAT GGAAGCCCTA A

21

## (2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATGGCCCCATT AAGTTTATCA C

21

## (2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AAAGCGTAGT CGGAAAGC

18

## (2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

ACCAATGATC TTTACGCAGA T

21

## (2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TAATCAGAGC GTACATCAGG A

21

## (2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

AGGGCATCAA TCAAGAATG

19

## (2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GAGACTTTTG AGGAGAAAGC A

21

## (2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GATCAAAAGA GCAAAAGGAG A

21

## (2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AACTGATGAA TACCTTCCCA G

21

## (2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TGATTAACCTT CTCCTTTGG T

21

## (2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TCGGCACAAC ATATGAGTTA C

21

## (2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CCCCCATAGA GAGAGATAGA G

21

## (2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AAGCACGGCC CAATAGAAT

19

## (2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

AGGATGTCCC TAGCTTTATT G

21

## (2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TCATTGCCCA AAGTGTTG

18

## (2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

CTCATCACCC CTCCAGAG

18

## (2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GATCCACGCC ATTTAAAC

18

## (2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGATACTCTG GTGCATGTTT

20

## (2) INFORMATION FOR SEQ ID NO:49:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GATCGCTCCG ATCTATCC

18

## (2) INFORMATION FOR SEQ ID NO:50:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGCGGCATCT ATGTTCTATG

20

## (2) INFORMATION FOR SEQ ID NO:51:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CCCAGTGCGA AGAGACTC

18

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ACACCTGCTC TGCACCAC

18

## (2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CTAACCCACG ACGACCCT

18

## (2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GCATGAGTGC ATGTGCAT

18

## (2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CTGCCACATG CTTTTCTG

18

## (2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CTGTAAAGAA GCTGGTCTGG A

21

## (2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

TTCTCCTCAT GGATGCGT

18

## (2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

CTATTTGGAA GTATGGGCTT CA

22

## (2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GAGGGCATCT ATGTGCAAC

19

## (2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GCTCAGAAGT TGC GTTTATG

20

## (2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

TTCCTTCACG TTTCTCTGTT AA

22

## (2) INFORMATION FOR SEQ ID NO:62:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CACATAAACC TAATGGGGTA CA

22

## (2) INFORMATION FOR SEQ ID NO:63:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

CCCAAAGGCG ATACCTATT

19

## (2) INFORMATION FOR SEQ ID NO:64:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

CCCACTTTCT CACTCTTTTC T

21

## (2) INFORMATION FOR SEQ ID NO:65:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GAGGTGAGTA CTATGCAAAT GC

22

## (2) INFORMATION FOR SEQ ID NO:66:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

CAGGCTTACC TAGCCTTCTC

20

## (2) INFORMATION FOR SEQ ID NO:67:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

CTATGGATGG CTGCTTGC

18

## (2) INFORMATION FOR SEQ ID NO:68:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GTCAGGCAGC AGAATGTG

18

## (2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

AAACCGTCCA GCGACTAC

18

## (2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GGAAGAACCA ATCCCATATC T

21

## (2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

AACATCCTGT CGGAAACAG

19

## (2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

TCATCACGTC TCTCTTTCAA C

21

## (2) INFORMATION FOR SEQ ID NO:73:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

TTGTGGCAGA ATCTCAAATT A

21

## (2) INFORMATION FOR SEQ ID NO:74:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CGACTGGTGA CATGTGAAG

19

## (2) INFORMATION FOR SEQ ID NO:75:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

AGTGAGGAAA GAATATGCTG G

21

## (2) INFORMATION FOR SEQ ID NO:76:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

TGGACTGAGA AACTGATTTG A

21

## (2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

CACAAATGTG AAGGTAAACA CT

22

## (2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

AATGGTACGG TTCAGGATG

19

## (2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

AGATGACGCA CGGAACAC

18

## (2) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

AGCATCATGT AGCAGGAGG

19

## (2) INFORMATION FOR SEQ ID NO:81:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

TTGCAGCACT GTCGTAGTC

19

## (2) INFORMATION FOR SEQ ID NO:82:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

GCGCGAGTGG AGTAGTAAG

19

## (2) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

AAGATTATGC AGATGAGACA CC

22

## (2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

GTTCCATGCT TTCCTTGG

18

## (2) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

CTCTGGGCTC TGTGTTAGAG T

21

## (2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

CTCTGGGCTC TGTGTTAGAG T

21

## (2) INFORMATION FOR SEQ ID NO:87:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

CTGCTTTCTC TGTTCACAGC

19

## (2) INFORMATION FOR SEQ ID NO:88:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

AATCGCTTAC TTGTAACCCA C

21

## (2) INFORMATION FOR SEQ ID NO:89:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

AAGAACGTAC GTCCCATAAA G

21

## (2) INFORMATION FOR SEQ ID NO:90:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CAAGGTAAAG TGACAAAGCA G

21

## (2) INFORMATION FOR SEQ ID NO:91:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

G TTCAGGATG AGGCGGAA

18

## (2) INFORMATION FOR SEQ ID NO:92:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

GTGATCATCG CAGGAGACC

19

## (2) INFORMATION FOR SEQ ID NO:93:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GGAGCCTGGA GTGAGAAC

18

## (2) INFORMATION FOR SEQ ID NO:94:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

CATGCTCACC TAACGTGG

18

## (2) INFORMATION FOR SEQ ID NO:95:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

ATCTGAACAC TTGAGCAACA A

21

## (2) INFORMATION FOR SEQ ID NO:96:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

ATAGACCGGA CCCATCAC

18

## (2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CGAACAGCGG GTACACCT

18

## (2) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GAGGTCAGCT TCCTCGATCT

20

## (2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GGAATCGTTC CTCCACAC

18

## (2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

CTTCCTCGGT GTCAGACG

18

## (2) INFORMATION FOR SEQ ID NO:101:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

ATGGAAACAT CAAAGTGGAT T

21

## (2) INFORMATION FOR SEQ ID NO:102:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

TGCTACCCTG ATGACCTGAT

20

## (2) INFORMATION FOR SEQ ID NO:103:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

ACCACTAGTC TCATATGAAG GG

22

## (2) INFORMATION FOR SEQ ID NO:104:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

GGTAGGTGGG TAGGGGTT

18

What is claimed:

1. A method of breeding for corn with increased kernel oil concentration comprising:

- 5           a) using one or more genetic markers to select a corn plant from a corn breeding population by marker-assisted selection, wherein the genetic markers are selected from the group consisting of s1375, s1384, s1394, s1416, s1422, s1432, s1457, s1480, s1476, s1478, s1484, s1500, s1513, s1529, s1544, s1545, s1630, s1633, s1647, s1750, s1756, s1757, s1767, s1772, s1774, s1780, s1797, 10           s1813, s1816, s1817, s1836, s1853, s1860, s1870, s1921, s1922, s1925, s1931, s1933, s1939, s1946, s1949, s2054, s2055, s2057, s2058, s2097, s2122, s2125, s2150, s2156 and s2175; and
- 15           b) crossing the selected corn plant with a second corn plant wherein the progeny corn plants of the cross display increased kernel oil concentration.

2. The method of claim 1 wherein the selected corn plant is member of an Alexho synthetic population or a progeny thereof.

3. A method for identifying corn plants or corn lines for use as parents for creation of a breeding population, the method comprising:

- 20           a) genotyping corn plants or corn lines with one or more genetic markers wherein the genetic markers are selected from the group consisting of s1375, s1384, s1394, s1416, s1422, s1432, s1457, s1480, s1476, s1478, s1484, s1500, s1513, s1529, s1544, s1545, s1630, s1633, s1647, s1750, s1756, s1757, s1767, s1772, s1774, 25           s1780, s1797, s1813, s1816, s1817, s1836, s1853, s1860, s1870, s1921, s1922, s1925, s1931, s1933, s1939, s1946, s1949, s2054, s2055, s2057, s2058, s2097, s2122, s2125, s2150, s2156 and s2175; and
- 30           b) identifying corn plants or corn lines which, based upon their genotype, are predicted to produce transgressive segregants for kernel oil concentration.

4. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s2054, s1647, s1500, s1545, s1774 and s2097.

35           5. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s1817 and s2057.

6. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s1860, s1931, s2150 and s1925.

7. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s1457, s2055, s1757, s2125, s1780, s1375, s1797, s1416, s1432 and s1921.
8. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s1544, s1633, s1384, s1813, s1767, s2058, s1933, s1513 and s1484.
9. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s1476, s1772, s1816, s2122 and s1836.
10. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s1939 and s1946.
11. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s1478, s1853 and s1949.
12. A trait locus controlling kernel oil concentration, the locus mapped by a genetic marker selected from the group consisting of s1630, s1422 and s2156.
13. A trait locus controlling kernel oil concentration, the locus mapped by the genetic marker s1756.
14. A trait locus controlling kernel oil concentration, the locus mapped by the genetic marker s1922.
15. A trait locus controlling kernel oil concentration, the locus mapped by the genetic marker s1529.
16. A trait locus controlling kernel oil concentration, the locus mapped by the genetic marker s1394.
17. A trait locus controlling kernel oil concentration, the locus mapped by the genetic marker s1750.
18. A trait locus controlling kernel oil concentration, the locus mapped by the genetic marker s1870.
19. A trait locus controlling kernel oil concentration, the locus mapped by the genetic marker s2175.
20. Corn plants that display increased kernel oil concentration produced by the method of Claim 1.

# INTERNATIONAL SEARCH REPORT

In national Application No

PCT/US 98/05550

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 476 524 A (LEON ALBERTO J ET AL) 19 December 1995 see esp. abstract and claims ---	1-20
A	KAHLER A.L. : "Association between enzyme marker loci and agronomic traits in maize" PROC. 40TH ANN. CORN AND SORGHUM RES. CONF. AM. SEED. TRADE ASSOC., - 1985 pages 66-85, XP002070561 see esp. p.78-82 --- -/-	1-20

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

7 July 1998

Date of mailing of the international search report

21/07/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040. Tx. 31 651 epo nl.  
Fax: (+31-70) 340-3016

Authorized officer

Müller, F

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/05550

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>GOLDMAN I. L. ET AL.,: "Molecular markers associated with amize kernel oil concentration in an Illinois high protein/ Illinois low protein cross" CROP SCI., vol. 34, - 1994 pages 908-915, XP002070562 see the whole document</p> <p>----</p>	1-20
A	<p>BERKE T. G. &amp; ROCHEFORD T.R.: "Quantitative trait loci for flowering, plant and ear height, and kernel traits in maize" CROP SCI., vol. 35, - 1995 pages 1542-1549, XP002070563 whole doc, see esp. pages 1545,1547,1548</p> <p>----</p>	1-20
A	<p>US 5 492 547 A (JOHNSON RICHARD) 20 February 1996 whole doc see esp claims and . e.g. column 5, line 55</p> <p>-----</p>	1-20

# INTERNATIONAL SEARCH REPORT

In. itional Application No

PCT/US 98/05550

Patent document  
cited in search report

Publication  
date

Patent family  
member(s)

Publication  
date

US 5476524 A 19-12-1995 NONE

US 5492547 A 20-02-1996 NONE